

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indihome merupakan salah satu penyedia layanan internet (ISP) yang ada di Indonesia yang jangkauan layanan sudah tercakup ke seluruh wilayah Indonesia. Sebuah penyedia layanan seperti Indihome pasti memiliki keluhan atau aduan tentang kepuasan atau kelayakan menggunakan Indihome, melalui data pengguna Twitter yang didapat, selanjutnya akan dilakukan analisis sentimen untuk mendapatkan akurasi nilai positif, negatif ataupun netral terhadap layanan Indihome.

Naïve Bayes Classifier adalah algoritma pembelajaran mesin probabilistik yang dapat digunakan untuk klasifikasi. Pada saat melakukan klasifikasi, algoritma kan melakukan pencarian nilai probabilitas yang tertinggi. Pada pemrosesan perhitungan Naïve Bayes hanya memerlukan tiga tahapan saja, dimulai dari perhitungan prior, likelihood dan posterior [20].

Menggunakan data *tweet* pada Twitter merupakan salah satu manfaat layanan yang disediakan untuk aduan atau keluhan terhadap Indihome. Dari data yang di ambil dari Twitter tersebut berupa ulasan keluhan atau aduan terhadap layanan untuk setiap orang memiliki perbedaan oleh karena itu Peneliti melakukan Penelitian ini melalui proses klasifikasi analisis sentimen pengguna Twitter menggunakan *Naive Bayes Classifier*. Pengambilan data *tweet* pada Twitter menggunakan script Bahasa Python kemudian data tersebut di klasifikasi menggunakan *Naive Bayes Classifier*. Proses nya meliputi hal berikut.

1. Mengumpulkan data *tweet* pada Twitter dengan beberapa kata kunci seperti, indihome, IndiHomeCare dsb.
2. Mengumpulkan data *tweet* menggunakan *script* bahasa Python dengan *package sncrape*
3. Melakukan pembobotan TF-IDF
4. Melakukan klasifikasi dan diolah untuk memperoleh nilai akurasi menggunakan *Naive Bayes Classifier*.
5. Melakukan Evaluasi dan Validasi.

Saat melakukan pengumpulan data *tweet* tersebut juga mengalami beberapa kendala. Kendala tersebut di antaranya.

1. Proses mengumpulkan data *tweet* memerlukan beberapa waktu.
2. Saat mengumpulkan data *tweet* terdapat duplikasi, mengandung kata tidak penting, seperti alamat laman, emoji, simbol.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan adalah bagaimana algoritma *Naive Bayes Classifier* mampu mengklasifikasi sentimen *tweet* terhadap layanan provider Indihome.

1.3 Batasan Masalah

Beberapa batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Pengumpulan data *tweet* menggunakan API Twitter pencarian dengan menggunakan kata kunci Indihome, IndiHomeCare.
2. Proses menganalisis sentimen algoritma *Naive Bayes Classifier* menggunakan bahasa python dan menggunakan package yang sudah tersedia yaitu *Snsrape & ScikitLearn*.
3. Data yang digunakan untuk proses klasifikasi menggunakan bahasa Indonesia.

1.4 Tujuan Penelitian

1. Tujuan dari penelitian ini dimaksudkan untuk mengetahui ketepatan akurasi klasifikasi algoritma *Naive Bayes Classifier* dengan metode *TF-IDF*.
2. Mengetahui hasil pandangan masyarakat pada klasifikasi data menggunakan algoritma *Naive Bayes Classifier* akan menghasilkan pemilih yang baik atau sebaliknya.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan bermanfaat untuk masyarakat Indonesia sebelum memutuskan untuk menggunakan layanan provider Indihome ataupun yang lainnya. Bagi peneliti selanjutnya agar dapat menjadi acuan ataupun pembanding dikemudian hari jikalau melakukan penelitian serupa.

1.6 Sistematika Penulisan

Sistematika Penulisan yang dapat Peneliti kemukakan atau jabarkan sebagai berikut

1. BAB I PENDAHULUAN

Pada bab ini peneliti menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Pada bab ini peneliti membahas lebih lanjut mengenai landasan teori atau tinjauan Pustaka tentang *Naïve Bayes Classifier*, *Preprocessing*, *TF-IDF*, *K-fold Cross Validation* dan *Confusion Matrix*. Teori yang telah dilakukan sebelumnya akan dijadikan sebagai dasar acuan penelitian ini.

3. BAB III METODE PENELITIAN

Pada bab ini peneliti menjabarkan analisis perancangan yang akan dibutuhkan pada sistem dan proses tahapan mengimplementasikan pada metode penelitian.

4. BAB IV HASIL DAN PEMBAHASAN

Pada bab ini peneliti mengimplementasikan penggunaan algoritma *Naïve Bayes Classifier* yang berisi hasil nilai akurasi pada pengujian data pengguna Twitter terkait layanan provider Indihome disajikan dalam bentuk web.

5. BAB V PENUTUP

Dalam bab ini berisi kesimpulan, saran dan harapan untuk peneliti selanjutnya, serta ringkasan dari hasil penelitian yang peneliti lakukan.